



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/815,591	03/23/2001	Victor Spivak	EMCCP074	5572
21912	7590	04/05/2006	EXAMINER	
VAN PELT, YI & JAMES LLP 10050 N. FOOTHILL BLVD #200 CUPERTINO, CA 95014				BASEHOAR, ADAM L
ART UNIT		PAPER NUMBER		
2178				

DATE MAILED: 04/05/2006

Please find below and/or attached an Office communication concerning this application or proceeding.

Office Action Summary	Application No.	Applicant(s)	
	09/815,591	SPIVAK ET AL.	
	Examiner	Art Unit	
	Adam L. Basehoar	2178	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --
Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1) Responsive to communication(s) filed on 26 January 2006.
 2a) This action is FINAL. 2b) This action is non-final.
 3) Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4) Claim(s) 1-20 is/are pending in the application.
 4a) Of the above claim(s) _____ is/are withdrawn from consideration.
 5) Claim(s) _____ is/are allowed.
 6) Claim(s) 1-20 is/are rejected.
 7) Claim(s) _____ is/are objected to.
 8) Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9) The specification is objected to by the Examiner.
 10) The drawing(s) filed on _____ is/are: a) accepted or b) objected to by the Examiner.
 Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).
 Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
 11) The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12) Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
 a) All b) Some * c) None of:
 1. Certified copies of the priority documents have been received.
 2. Certified copies of the priority documents have been received in Application No. _____.
 3. Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) <input checked="" type="checkbox"/> Notice of References Cited (PTO-892)	4) <input type="checkbox"/> Interview Summary (PTO-413)
2) <input type="checkbox"/> Notice of Draftsperson's Patent Drawing Review (PTO-948)	Paper No(s)/Mail Date. _____
3) <input type="checkbox"/> Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08) Paper No(s)/Mail Date _____.	5) <input type="checkbox"/> Notice of Informal Patent Application (PTO-152)
	6) <input type="checkbox"/> Other: _____.

DETAILED ACTION

1. This action is responsive to communications: The RCE filed 01/26/06.
2. The rejection of claims 1-5 and 7-20 under 35 U.S.C. 103(a) as being unpatentable over Russell-Falla et al (US: 6,675,162 01/06/04) in view of Weiser et al (US-5,982,507 11/09/99) have been withdrawn as necessitated by Amendment.
3. The rejection of claim 6 under 35 U.S.C. 103(a) as being unpatentable over Russell-Falla et al (US: 6,675,162 01/06/04) in view of Weiser et al (US-5,982,507 11/09/99) in further view of W3C's, "Extensible Markup Language (XML) 1.0", 02/10/98, pp. 1-2, <http://www.w3.org/TR/1998/REC-xml-19980210> has been withdrawn as necessitated by Amendment.
4. Claims 1-20 are pending in the case. Claims 1, 7, and 14 are independent claims.

Claim Rejections - 35 USC § 103

5. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.
6. Claims 1-5 and 7-20 are rejected under 35 U.S.C. 103(a) as being unpatentable over Russell-Falla et al (US: 6,675,162 01/06/04) in view of Weiser et al (US-5,982,507 11/09/99) in further view of Doerre et al (US-6,446,061 09/03/02).

-In regards to independent claims 1, 7, and 14, Russell-Falla teaches a computer-implemented method comprising a processor (Abstract) and memory (Fig. 1: 30) connected to said processor, wherein the method further comprises;

recognizing a concept (column 2, lines 54-63) that represents a basic idea (content category)(column 2, lines 35-39; column 4, lines 32-47) in a document format (column 2, lines 35-39; column 3, lines 17-20);

incorporating said concept in a concept model (i.e. “pornographic”, “commercial solicitations”, “racist”, “good”, “bad”, etc)(column 3, lines 39-43 & 60-67; column 8, lines 20-48); and

using said conceptual model (column 3, lines 39-67) to determine whether said document was responsive to a search query (column 2, lines 5-22: “search engine”; column 3, lines 2-7; column 5, lines 53-67; column 6, lines 1-7)(i.e. the conceptual model determines where the document was responsive to a search query based on whether the categorized document was blocked or viewed by the user).

Russell-Falla further teaches wherein the document format could be any number of common document formats including an electronic email message, a word processing document, hypertext document, and any number of other types of documents (columns 3 & 4, lines 23-26 & 51-53). Russell-Falla does not teach wherein the initial document format have to be converted to one of the common document formats to be processed. Weiser et al teach converting a document format (email message) from an email format to a common generic format (column 12, lines 53-55). It would have been obvious to one of ordinary skill in the art at the time of the invention for Russell-Falla to have converted its initial format document to one of the common document

formats listed above, because Weiser et al teach by doing so the common format can be understandable by the document system (column 12, lines 44-56)(i.e. converting document to a format able to be processed by the a specific system provides the obvious advantage of being able to process the document in that system).

Russell-Falla does not specifically teach wherein the concept in the concept model was associated with a hierarchical conceptual taxonomy specifying at least one relationship between two or more concepts. Doerre et al teach the task of clustering documents by associating a concept (column 16, lines 21-23) in a concept model with a hierarchical conceptual taxonomy (column 4, line 39: “a tree-structured taxonomy-hierarchy”) specifying relationships between two or more concepts (column 4, lines 40-52 & 66-67; column 5, lines 1-6; column 12, lines 17-34; column 52-67). It would have been obvious to one of ordinary skill in the art at the time of the invention at the time of the invention for the concept in the concept model of Russell-Falla to have been associated with a hierarchical conceptual taxonomy, because Doerre et al teach that associating concepts (column 16, lines 21-23) with a taxonomy hierarchy provide the benefit of good coherence and selectivity of the taxonomy as well as provides good orientation to a user traversing the taxonomy (column 5, lines 33-43). Thus Russell-Falla would gain the benefit of being able to determine the similarities/differences of the categorized pages in greater detail than just “bad” or “good” based on the defined threshold.

-In regard to dependent claims 2 and 8, Russell-Falla teaches identifying a plurality of features (column 4, lines 59-61: “identify the regular expressions”) in said document format,

wherein said plurality of features represent evidence (“useful in discriminating a specific category of information”)(column 4, lines 61-66) of said concept in said format.

-In regard to dependent claims 3 and 9, Russell-Falla teach calculating a concept weight for said concept (“calculating a rating of the page”)(column 3, lines 54-57) using a plurality of feature weights (“requires a weighting be provided for each word of phrase in the list”)(column 3, lines 46-57) associated with said plurality of features (“regular expressions”)(column 2, lines 55-59; column 8, lines 9-19) wherein said concept weight represents a recognition confidence level for said concept (column 3, lines 54-57);

comparing said concept weight with a predetermined thresholds (column 2, lines 64-67; column 3, lines 1-16).

-In regard to dependent claims 4, 11, 13, and 19, Russell-Falla teaches by referencing said concept model (content category)(column 2, lines 35-39), generating an auto-attribute/category (column 8, lines 39-51), said auto-attribute/category being a descriptive label (i.e. “pornographic”, “commercial solicitations”, “racist”, “good”, “bad”, etc)(column 3, lines 39-43 & 60-67; column 8, lines 43-45) for said format/category document.

-In regard to dependent claims 5, 12, 18, and 20, Russell-Falla teaches by referencing said concept model (content category)(column 2, lines 35-39), assigning said document format to a subject category/modeling directory (i.e. “pornographic”, “commercial solicitations”, “racist”, “good”, “bad”, etc)(column 3, lines 39-43 & 60-67; column 8, lines 43-45) in a categorization

taxonomy (column 4, lines 34-45) including a plurality of categories (i.e. "pornographic", "commercial solicitations", "racist", "good", "bad", etc)(column 3, lines 39-43 & 60-67; column 8, lines 43-45).

-In regard to dependent claim 10, Russell-Falla teaches incorporating said recognition confidence level (category threshold) (column 2, lines 64-67; column 3, lines 1-16) in said conceptual model (content category)(column 2, lines 35-39) based on the training data (column 6, lines 52-67; column 7, lines 1-67).

-In regard to dependent claim 15, as shown above, Russell-Falla teaches wherein the common document format was hypertext (HTML) web pages (column 1, lines 33-37)(Fig. 1: 12) or other like information content (column 3, lines 17-22; column 6, lines 25-28; column 8, lines 20-61: "file directories", "email messages", "database records", "other web pages", etc). Russell-Falla does not teach wherein the initial document format have to be converted to one of the common document formats to be processed. Weiser et al teach converting a document format (email message) from an email format to a common generic format (column 12, lines 53-55). It would have been obvious to one of ordinary skill in the art at the time of the invention for Russell-Falla to have converted its initial format document to one of the common document formats listed above, because Weiser et al teach by doing so the common format can be understandable by the document system (column 12, lines 44-56)(i.e. converting document to a format able to be processed by the specific system provides the obvious advantage of being able to process the document in that system).

-In regard to dependent claim 16, Russell-Falla teaches separating the text content from said initial format document for categorizing documents based on statistical techniques (column 2, lines 52-59). As shown above in dependent claim 15, Russell-Falla does not teach converting the initial document format into a common document format. Weiser et al teach converting a document format (email message) from an email format to a common generic format (column 12, lines 53-55). It would have been obvious to one of ordinary skill in the art at the time of the invention for Russell-Falla to have converted its initial format document to one of the common document formats listed above, because Weiser et al teach by doing so the common format can be understandable by the document system (column 12, lines 44-56)(i.e. converting document to a format able to be processed by the specific system provides the obvious advantage of being able to process the document in that system).

wherein it would have also been obvious to incorporate the text from the initial document into the said common document, because Russell-Falla teaches the textual content was what was needed to categorize the incoming documents (column 4, lines 57-66).

-In regard to dependent claim 17, Russell-Falla teaches identifying a plurality of features (column 4, lines 59-61: “identify the regular expressions”) in said document format, wherein said plurality of features represent evidence (“useful in discriminating a specific category of information”)(column 4, lines 61-66) of said concept in said format. Russell-Falla further teaches calculating a concept weight for said concept (“calculating a rating of the page”)(column 3, lines 54-57) using a plurality of feature weights (“requires a weighting be provided for each word of

phrase in the list”)(column 3, lines 46-57) associated with said plurality of features (“regular expressions”)(column 2, lines 55-59; column 8, lines 9-19), wherein said concept weight represents a recognition confidence level for said concept (column 3, lines 54-57); and comparing said concept weight with a predetermined thresholds (column 2, lines 64-67; column 3, lines 1-16).

7. Claim 6 is rejected under 35 U.S.C. 103(a) as being unpatentable over Russell-Falla et al (US: 6,675,162 01/06/04) in view of Weiser et al (US-5,982,507 11/09/99) in further view of Doerre et al (US-6,446,061 09/03/02) in further view of W3C’s, “Extensible Markup Language (XML) 1.0”, 02/10/98, pp. 1-2, <http://www.w3.org/TR/1998/REC-xml-19980210>.

-In regard to dependent claim 6, Russell-Falla teach wherein a common document format was hypertext (HTML) web pages (column 1, lines 33-37)(Fig. 1: 12) or other like information content (column 3, lines 17-22; column 6, lines 25-28; column 8, lines 20-61: “file directories”, “email messages”, “database records”, “other web pages”, etc). Russell-Falla does not specifically teach wherein a common format was an XML document. W3C teaches wherein using XML was notoriously well known in the art for web applications (pp. 2: Section 1.1). It would have been obvious to one of ordinary skill in the art at the time of the invention, for one of the common formats of Russell-Falla to have been XML, because W3C teaches that the XML format provides the benefits of being easy to create, being easy to write programs which process XML documents, and being human-legible and reasonably clear (pp. 2: Section 1.1). It was also notoriously well known in the art at the time of the invention that XML was an International

document standard and well known for its separation of data content which was the main embodiment of the Russell-Falla reference (column 4, lines 59-66; column 8, lines 20-38).

Response to Arguments

8. Applicant's arguments with respect to claim 1 have been considered but are moot in view of the new ground(s) of rejection.

Conclusion

9. The prior art made of record and not relied upon is considered pertinent to applicant's disclosure.

US-6,185,550	02-2001	Snow et al.
US-6,990,628	01-2006	Palmer et al.
US-6,389,436	05-2002	Charkrabarti et al.

10. Any inquiry concerning this communication or earlier communications from the examiner should be directed to Adam L. Basehoar whose telephone number is (571)-272-4121. The examiner can normally be reached on M-F: 7:00am - 4:00pm.

Art Unit: 2178

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Steve Hong can be reached on (571) 272-4124. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

ALB



STEPHEN HONG
SUPERVISORY PATENT EXAMINER